As Chairman Pressler's draft legislation recognizes, the electromagnetic spectrum is a valuable and increasingly scarce resource that should be managed in a way that maximizes opportunities for technological advancements. The development of new services that efficiently use spectrum should not be impeded by regulatory restrictions on spectrum use that promote relatively inefficient, less advanced services.

Given the limited supply of usable spectrum, tough decisions inevitably have to be made regarding its best uses. As a general matter, members of CICATS believe that the marketplace, not government, is best equipped to make these decisions. Government policies should be tailored to protecting the public interest in the most desirable uses of spectrum, but the public should be the final arbiter of which uses best serve its interests.

If the process for allocating spectrum is slow or cumbersome, or if spectrum regulation is unduly restrictive, development of new spectrum-based technologies will be discouraged. Whether or not Congress determines that spectrum should be auctioned, government policies should aim to ensure that spectrum is available when emerging advanced services require it. Any regulation of spectrum use that hampers technological progress should be unequivocally justified by clear, compelling benefits to the public which could not be achieved absent that regulation.

For example, restrictions on interference with other uses of spectrum, and regulations designed to ensure adequate spectrum for public safety, transportation, and national security uses clearly benefit the public and are therefore generally justifiable. In contrast, the public interest would be poorly served by adoption of a standard for spectrum use that would impose significant costs on consumers and discourage future technological development.

Mandating the digital broadcast television standard (DTV) proposed by the Advanced Television Systems Committee (ATSC) will have both of these negative effects. It is costly because the standard is not layered. All receivers must be capable of decoding the highest resolution transmissions regardless of whether they are capable of displaying that resolution. Making the standard a law will lock in today's view of technological capability for a very long time. Any modifications or improvements will have to run the gauntlet of a long and arduous government approval process, something with which even the members of ATSC are already too familiar.

We do not mean to diminish the hard work of the ATSC The standard they have proposed contains some noteworthy attributes, many of which the computer industry supports. And if proponents of that standard believe it will best serve the public's needs and tastes, they should be free to produce and market products meeting the standard.

But those of us who think we can build a better mousetrap -- or digital TV receiver -- should be permitted, in fact, encouraged, to try. We should not be forced to overcome a government-mandated competitive advantage, which adoption of the standard would amount to for its advocates. The public should be allowed to decide what's best for them. Isn't that what drives a free market economy and results in the greatest economic efficiency?

The robustness of this country's computer and software industries is proof that great efficiency, innovation, and productivity can be achieved quickly when industry standards are *voluntarily* set in response to demand. Voluntary standards work. Look at cellular telephones. The FCC recognized that the detailed standards it originally prescribed for cellular telephony were holding back technological development in that industry, and it decided to relax its standards and let the industry establish more advanced standards with minimal government oversight. In doing so, the Commission acknowledged that too much government-specification of industry standards can inhibit technological progress and the availability to consumers of improved services. With Personal Communications Service, or "PCS," the FCC took an even more liberal industry-based approach to standards-setting. It should do the same with digital TV.

Our domestic computer and software industries -- like many other industries -- have thrived in large measure because of two factors: a minimum of government regulation, and open system architecture that permits hardware and software produced by many different firms to interconnect smoothly and encourages rapid, market-driven innovation. Both of these factors would be negated by the FCC's adoption of the Grand Alliance DTV standard, and the public would pay the price.

Let's look for a moment at that standard. Beyond public policy and macroeconomic, free-market considerations, there are both consumer interests and technical drawbacks that make adoption of the standard bad policy.

First, the standard does not provide for a way to manufacture low cost receivers. The encoding technique is monolithic. If a broadcaster chooses to send the highest resolution format a receiver must include all of the circuitry necessary to decode that format. In a layered system, two signals are sent in the channel simultaneously. A low resolution, easily decodable version for smaller cheaper receivers and a higher resolution detail enhancement signal for use by larger, more expensive high definition receivers. In the ATSC system, all receivers, even a little 2" portable must be burdened with means to decode resolution only perceivable on a large screen home theater unit. We have determined that even five years from now a full ATSC decoder will be three times the cost of a base layer decoder. Using the ATSC system will drive up the cost of smaller devices and require consumers to pay for capabilities they may neither need not want.

Second, from a technical perspective, the Grand Alliance standard is a poor compromise, particularly with respect to its video formats. The standard incorporates an obsolete technology, interlaced scanning, that produces an inferior picture and makes inter-conversion for computer uses difficult. In fact, ABC recently announced at a meeting of its affiliates that the network is leaning heavily toward the use of progressive scanning for all its high-definition TV production, because progressive scanning produces a better picture and is less expensive. Even ACATS has admitted that progressive scanning is better. Interlace was an appropriate scheme for the analog television of 40 years ago, but it has no place in a modern digital compressed transmission system.

But broadcasters have been using interlaced scanning for over 40 years. Despite what ABC has said, local stations will have little incentive to replace it with progressive scanning if the FCC adopts a digital standard that allows them to continue to use interlaced. And this is a critical issue for the computer industry because interlaced scanning is unacceptable for text and other computer applications. Any interlaced transmission would have to be converted at the receiver if it is to be used with a computer application. Again, added costs for the consumers.

These limitations of the ATSC proposal would make it more expensive for the domestic computer and software industries to create products -- both hardware and software -- that could enhance the usefulness of digital TVs by marrying digital broadcasting and computers. For these reasons, when ACATS voted to recommend the ATSC standard to the FCC, I abstained.

NTSC broadcast television is transmitted in an analog format. Computer data is digital. As long as analog broadcasting continues, the convergence of TVs and computers will be delayed. But with the advent of digital TV, interactive applications, multimedia, and data sharing between TV and computers are all possible. The products and services that data sharing could make possible are limitless. Microsoft and other firms have committed hundreds of millions of dollars to research and development of products and services that combine computers and TVs; but these products may never reach the stores, at least not at affordable prices, if overly detailed and restrictive regulatory requirements obstruct full compatibility, product development, and competition.

The Grand Alliance says that its proposal provides "adequate" compatibility with computers. We disagree. True, some of the 18 video formats are consistent with computer applications, but the standard also includes a number of inconsistent formats. And if a mandated standard incorporates even one computer-unfriendly format, receiving equipment will need to perform additional conversion and decoding of transmissions to enable interaction with computer applications, the added cost of which will fall on the consumer.

Why does the computer industry care about these issues? Two reasons, mainly. First, we don't want government regulation to freeze technological development without a compelling justification. We think a better DTV standard is possible, and we want the freedom to try it out on the market. Second, our industry knows that computers and TVs can, and will, converge, and we want to be able to develop products that take advantage of that convergence and bring new benefits to the public. Who knows how advanced our National Information Infrastructure can become, if it is allowed to

In short, in this case, we think voluntary industry standards are better for everyone than government-mandated standards. We understand the value of minimal government-sanctioned technical transmission standards for digital broadcasting, including standards for low level digital bitstream format and modulation technique to prevent interference with other services and would not object to adoption of the ATSC's proposals with respect to those parameters, absent any specified video format.

But specifying a video format is unnecessary and potentially problematic -- exponentially so with 18 formats. We think the marketplace should dictate what video formats it wants. But if the Congress and the FCC find that the public interest would be served by the FCC's adoption of a standard video format for digital television, the standard it adopts should be the best possible. That would not include the hodgepodge of 18 different video formats the FCC is currently considering. If a standard is to be adopted at all, CICATS would propose a simpler, more technologically advanced minimum standard, offering wider compatibility and more flexibility to develop enhancements, if the marketplace warrants.

A year ago, computing capability was not sufficient for the level of convergence of TVs and computers and the sophistication of applications we are developing. It is now. Largely because computer technology is advancing at an exponential rate, the computer industry's interest in advanced television emerged relatively recently. The question should not be *whether* TVs and computers will ever converge seamlessly, but *when* and whether it will be affordable. If the FCC adopts the proposed ATSC standard, the "when" will be years from now -- some say 5 to 7 years later than if the Commission adopts a simpler standard or no standard at all. And when convergence finally arrives, the average consumer will be hard-pressed to afford the advanced products and services convergence will spawn if government regulation imposes a cumbersome, overly complex DTV standard.

If the price of digital receivers and decoders is unnecessarily inflated, the day stations will migrate to all-digital broadcasting will be delayed, and so, in turn, will the day analog spectrum is freed for new uses. In the meantime, precious spectrum could be wasted and consumers could be deprived of better, and cheaper, products and services.

Thank you for your time. I would be pleased to answer any questions you might have.

6

Advanced Television Systems for Terrestrial Broadcasting: Some Problems and Some Proposed Solutions

WILLIAM F. SCHREIBER, FELLOW, IEEE

Invited Paper

The first part of this paper discusses the requirements that must be met by a new television broadcasting system to maximize its acceptability to the various stakeholders, including broadcasters, equipment manufacturers, program producers, regulatory authorities, and viewers. The most important performance factors are efficient use of over-the-air spectrum, coverage versus quality, cost, interoperability, and the existence of a practical transition scenario. It is concluded that all receivers need not have the same performance, and that low-cost receivers must be available for noncritical locations in the home. If this variation in price and performance is made possible by appropriate system design, then interoperability is facilitated and nondisruptive improvement over time is made possible, as desired by the Federal Communications Commission.

In the second part of the paper, techniques that may permit meeting these requirements are discussed. These include joint multiresolution source and channel coding, multicarrier modulation, and hybrid analog/digital coding and transmission. The analog transform coefficients are subjected to spread-spectrum processing, and coded orthogonal frequency-division multiplex (COFDM) is applied to the complex hybrid symbols to be transmitted through the channel. Various methods of equalization and of improving noise, interference, and multipath rejection are compared. Finally, an example is given of a system that meets the various requirements by making use of a number of the techniques discussed. The system provides extended coverage, albeit at lower quality than currently proposed all-digital systems, and equal or higher quality than such systems in much of their service area. It also features self-optimization at each receiver, depending on signal quality and receiver characteristics, and facilitates the design of receivers of lower cost and performance for less-critical applications.

I. Introduction

Since the proposal by General Instrument Corporation (GI) in 1990 for all-digital terrestrial broadcasting of high-definition television (HDTV), remarkable enthusiasm has developed in many quarters for what is, in reality, a truly radical departure from current practice. Digital technology, of course, had been widely accepted in many fields, including television post-production and video recording. Digital compression had been the subject of an international standardization process for several years under the aegis of JPEG and MPEG. The most notable features of the GI proposal were the degree of compression employed and the use of digital transmission technology. All of the earlier HDTV proposals, without exception, had made use of digital signal processing at encoder and decoder and had used some degree of digital compression. None, however, had used digital transmission. That technique, to the best of the author's knowledge, is currently employed in no terrestrial broadcasting system except for JTIDS, a US military system based on spread spectrum. The main applications of digital transmission are currently in wired point-to-point systems and in satellite communications. In those media, channel impairments are much less severe and receiver CNR[1] is much more uniform than found in terrestrial broadcasting. There, noise, interference, and multipath are particularly troublesome, and CNR varies enormously over the population of receivers.

For these and other reasons, many in the TV industry had thought that all-digital systems were very far in the future. Digital proposals had often been viewed as roundabout efforts to delay HDTV. Likewise, it had been the generally held (but incorrect) view that *any* amount of compression would be unacceptable because of loss of quality.

This being the case, it is natural to wonder what was the primary motivation for using digital transmission. A number of reasons were often given—better utilization of channel capacity, suppression of multipath effects, and higher resistance to noise and interference. Among those in the computer community who have been pressing for easy interoperability between the TV broadcasting format and

Manuscript received July 1, 1994; revised December 22, 1994. This work was supported in part by Scitex America, Inc.

The author is with The Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

IEEE Log Number 9410562.

[1] In this paper, CNR is used for the signal-to-noise ratio at the receiver terminals, and SNR is used when referring to the recovered video.

ormats useful for displaying video on computer screens. It is often averred that digital transmission enhances interoperability. *All of these reasons are fallacious.*

As both system proponents and the Advisory Committee on Advanced Television Systems (ACATS)[2] personnel got more deeply into the details of the all-digital proposals, the first three alleged advantages were heard less and less. The interoperability argument, however, is still voiced. Since this issue is central to the subject matter of this paper, it is dealt with in some detail in Section III-A-3. The other matters are considered briefly in the Appendix. What we shall see is that digital transmission generally makes less efficient use of channel capacity than analog or hybrid analog/digital transmission. However, the very high compression ratio (50–80) achieved by the currently proposed HDTV systems reduces the data rate sufficiently so that coded HDTV signals can be transmitted at a gross data rate of 20–25 Mb/s, which, under the right conditions, can be transmitted in the usual 6-MHz channel. The real question is whether all-digital transmission is *required* in order to attain the required high levels of compression in the source coder. As we shall show later, hybrid transmission also permits effective compression.

In the earlier American TV standardization processes (1941 and 1953), a vigorous consumer-electronics industry spearheaded by RCA did the development work and the Federal Communications Commission (FCC) adopted, for the most part, the transmission format recommended by the industry. However, by the time the formal HDTV standards-setting process started in 1987, the US consumer-electronics industry had been decimated and proposals for federal funding were subsequently rejected. Thus the various development projects have been grossly underfunded and all competitors have worked under unrealistically short time schedules.[3] As a result, even though the development work has been of remarkably high quality, many issues were not given sufficient study. In particular, not enough attention was directed toward the characteristics that an entirely new TV system ought to have. Equally important, very little attention was given to coding methods for the terrestrial channel until after GI made its proposal. To this date, work on channel coding in the US remains far behind that in Europe. These topics are the main subject of this paper.

When speaking of "currently proposed" HDTV systems, we are referring to the Grand Alliance (GA) scheme, [1] which is a melding of the four all-digital systems that were tested by the Advanced Television Test Center (ATTC). Many of the features of the "ideal" system discussed below are intended to deal specifically with aspects of the GA system that the author feels are questionable for terrestrial broadcasting.

The paper is primarily addressed to HDTV in the US. The situation in Europe is quite different, for a number of reasons. In Europe, as compared with the US, government entities play a much larger role, the domestic consumer-electronics industry is much stronger, cable is less widespread and evidently of higher technical quality, and satellite broadcasting is further advanced. Many fewer terrestrial channels are available to each viewer, and a considerable investment was made in HD-MAC, a failed system. There has been almost no controversy over interlace, as the path to digital broadcasting seems to have been laid out in the expectation of very few changes in the studio. Digital television of standard definition is the evident current intention of cable and satellite interests in the US. In Europe, this also seems to be the case. In both areas, those planning digital services are all saying something about eventually going to HDTV, but ensuring that the first digital receivers can still function seems not to be getting much attention.

Many of the issues addressed in this paper involve political or economic considerations as well as technical matters. Therefore, the analysis cannot be entirely objective, nor can it always be quantitative. New television systems can no more be designed completely on a quantitative basis than can automobiles. Qualitative analysis, for example on the question of the best use of spectrum, is the only way to deal with some very important matters. It should be clear from the context which statements in the paper are the author's opinion and which are based on quantitative analysis.

II. Problems of Television Broadcasting

A. Performance Factors in Terrestrial Broadcasting

On the reasonable assumption that good solutions are most likely to be found when the problems are most completely and accurately defined, we shall now set forth the desirable properties of an entirely new TV system. Note that this is a much more difficult task than that encountered in typical new product development. A TV system must not only produce profits for a company; it must serve the public interest for many years to come and it must be acceptable to the many stakeholders—broadcasters, program producers, equipment manufacturers, and the viewing public. In the case of HDTV, an even wider constituency has emerged with the increasing use of video in other fields such as the computer industry and military equipment, and the often-expressed desire for interoperability among the various applications.[4]

1) Spectrum Efficiency: Standing at the head of any list of desirable attributes of a terrestrial broadcasting system is the effective use of radio spectrum. A useful figure of

[2] ACATS was appointed by the FCC in 1987 to conduct the inquiry that is leading to the promulgation of HDTV terrestrial broadcasting standards.

[3] It is not clear that the tight schedules have produced a quicker result. The reverse may be true, since the optimistic schedules have never been met. In addition, the intensity at which the work was carried out (one team worked on Christmas Day!) precluded much consideration of alternative technologies.

[4] This paper does not concern itself with issues, real as they are, such as the importance of electronic imaging to the economic security of the US, and the possibility that an entirely new development such as HDTV might be a way for the country to revive its moribund consumer electronics industry [2].

merit, which we shall call spectrum efficiency[5] in this paper, is defined as the number of different programs of a certain technical picture and sound quality that are made available to each viewer per unit of allocated spectrum. This measure depends both on the quality that can be delivered with a fixed bandwidth per program and the number of different programs that can be delivered within the overall spectrum allocation. These properties are associated with source coding and channel coding, respectively. It is obvious that source coding is concerned with data compression, while channel coding is concerned with interference performance. The two are of equal value and importance. They are further discussed in Section II-B.

The overwhelming significance of the efficient use of spectrum arises from the fact that there is considerably more demand than supply. The FCC, required by the Communications Act to regulate in the "public interest, convenience, and necessity," must constantly adjudicate among the claims of various parties for spectrum assignments. As mobile applications have become much more common, this has become an increasingly difficult job. Television is at the root of the problem since it has more than 400 MHz of the most easily used spectrum. A highly desirable outcome of the HDTV standard-setting process would be to maintain or even increase the present level of service while substantially decreasing the total allocated bandwidth.

2) Coverage versus Quality: Commercial broadcasters, who derive their income from advertising, live or die according to their coverage, since they get paid on a per-viewer basis. The main way in which they compete with each other is by means of program popularity, but they must reach the viewer in order to compete. They are therefore most reluctant to accept any new system that significantly reduces coverage. Unfortunately, coverage must be traded off against technical quality, since the latter depends on the information rate to the receiver. The theoretically maximum information rate per unit bandwidth depends primarily on the signal-to-noise and signal-to-interference ratios at the receiver. The higher the CNR required for a given quality, the smaller the coverage, whether limited by noise or by interference. This tradeoff is also affected by the compression achieved in the source coder, as compression decreases the information rate needed for a given quality. Thus the fundamental question in coverage is whether sufficient compression can be achieved in the source coder to maintain coverage with a given quality while at the same time permitting a practical transition scenario from today's National Television Systems Committee (NTSC)[6] broadcasting to whatever will replace it. Because it has such low spectrum efficiency, almost everyone now agrees, albeit reluctantly, that NTSC must eventually be replaced.

a) Noise performance: The theoretical (Shannon) capacity, in bits per second, that is available to a receiver connected to an analog channel is proportional to the product of bandwidth and (1 + CNR) in dB. When the input is a multilevel signal, so as to effect digital transmission in such a channel, the error-free recovered data rate is usually less than the Shannon rate for a number of reasons. Clearly, if the level-spacing is too large relative to the RMS noise, the input must have a data rate less than the Shannon rate. No kind of postprocessing can cure this problem. If the level spacing is fine enough so as not to reduce the input data rate excessively, error correction must be used.[7] Very effective error-correction methods, using trellis coding and Viterbi decoding, are now available. Even so, the net recovered data rate, R, is reduced by any remaining errors according to the relationship

$$R = R_o - H(e)$$

where R_o is the error-free transmission rate, i.e., the maximum possible entropy of such a multilevel input signal, and $H(e)$ is the equivocation, or entropy of the error distribution. Essentially, the data throughput rate is reduced by the amount of information required to identify (and correct) the errors [3].

When high compression ratios are achieved in the source coder, the recovered information is usually more readily damaged by transmission errors. Thus, error correction must be used. Shannon proved that codes exist that permit transmission as close to the theoretical rate as desired with as small a bit error rate (BER) as desired. This involves removing all of the redundancy from the transmitted signal. If we could do that, we would find that the signal was very fragile and that it took a long time to resynchronize after an error. High channel-coding efficiency also implies a large amount of delay and more expensive processing. In practice, it is unusual to achieve even 75% of the Shannon rate, even at the given threshold CNR. In broadcasting, most of the receivers have a higher CNR than that at threshold. At these sites, channel capacity is higher than the transmission rate and, therefore, the efficiency is lower.

Another characteristic of effective error-correction systems is a very sharp threshold. In a heavily coded system, less than a 1-dB change in CNR takes one from perfect reception to no reception at all. This so-called "cliff effect" is not entirely a bad thing. In order to minimize the no-man's land between two different stations on the same channel, a sharp threshold may be helpful. However, it also leads to performance that is very different in character near the boundary of service from what is achieved in analog transmission. The viewing public is used to pictures getting a little worse or a little better, but not disappearing completely, every time a truck goes by or the character of

[5] Unfortunately, this term is sometimes used with the more limited meaning of transmission rate in bits per cycle of bandwidth.

[6] NTSC, an industry group, promulgated standards for television broadcasting in the US in 1941 and 1953. The proposed standards were adopted with little change by the FCC.

[7] Although this discussion is in terms of quantization of single samples, it applies equally to more sophisticated schemes in which a long train of samples is coded together as a single message. The selection of a finite number of such possible messages from the infinite number that is associated with unquantized analog samples is equivalent, for this argument, to the quantization mentioned above. The decision at the receiver as to which message was transmitted on the basis of minimum distance in multidimensional signal space is equivalent to the selection, at the receiver, of the nearest level to the received sample value.

on interfering signal changes. Although these considerations are very important, they are not amenable to quantitative analysis.

b) Co-channel interference: While noise can be effectively suppressed by raising the signal power, this increases interference to nearby stations. If all stations raised their power by the same amount, noise sensitivity would go down, but the interference situation would be unchanged. In the transition scenario in which HDTV and NTSC are to coexist for 15 years, the HDTV stations will be limited in power so as not to reduce the coverage of NTSC stations significantly. As a result, they may be noise-limited in portions of their intended coverage areas where there is no potential interference from an existing NTSC station.

One of the main defects of NTSC is that all transmissions are highly correlated. This causes one picture to appear on top of another when there is interference.[8] For a given strength signal interfering with an analog video transmission, the least-perceptible effect is produced by signals that appear to be *random* noise. Wisely, this has been done in HDTV, where each signal appears to be random noise to other signals. This means that the required signal/interference ratio is virtually identical to the required signal/noise ratio.

c) Adjacent-channel interference: This is a different question from cochannel interference, since there seems to be no reason why we cannot use adjacent channels in the same area provided that receivers have good-enough selectivity. The problem arises when a viewer tries to receive a distant station when there is a nearby station in an adjacent channel. This is not only a question of selectivity, it is also a question of out-of-band radiation by the nearby station. There is a limit to how much attenuation can be provided by filters at the transmitter without unduly distorting the in-band signal.

This problem can be solved either by placing all transmitters in any one city at the same location,[9] or by making use of modulation methods that inherently restrict out-of-band radiation, as in OFDM. On cable, where all signals are of the same amplitude, typical receivers have no trouble discriminating against signals in the adjacent channel.

d) Multipath: The final obstacle to effective use of the terrestrial transmission channel is multipath, i.e., the reception of a number of signals that have traveled over different paths from transmitting antenna to receiving antenna and therefore arrive displaced in time. In analog systems this causes the familiar ghosts, while in digital systems, it raises the error rate. The effect in digital systems is so strong that multipath must be essentially eliminated in order to permit any useful transmission at all. Elimination of ghosts in analog systems greatly improves picture quality, but the presence of ghosts does not generally make the service completely unusable.

Multipath is a linear distortion, so the effect is to produce a nonuniform frequency response across the channel, exactly as if an unwanted linear filter were processing the transmitted signal. It therefore can be corrected, within limits, by the use of the appropriate compensating filter, a process called linear equalization. First used in telephone circuits, the theory and practice of linear equalization are highly developed [4]. In the presence of noise, there are limits on what can be done. Large echoes cause deep notches in the frequency response, and correction by linear equalization may greatly increase the noise level. Noise in the received signal also makes determination of the parameters of the equalization filter slower and more difficult. For all these reasons, effective equalization requires a lot of computation. For example, in the GI system, one-third of the receiver signal-processing circuitry is used for this function [5].

3) Cost to the Stakeholders: In order for a new TV system to go on the air, it must be accepted by broadcasters, equipment manufacturers, and program producers. Once these difficult hurdles are surmounted, final success depends on acceptance by advertisers and viewers, who, in the end, will pay for the entire system. The different stakeholders have different needs [6], but near the top of everyone's list is cost.

a) Broadcasters: As mentioned above, broadcasters have little motivation to shift to HDTV except to help preserve audience share. If it appears that there is no way to stay in business while avoiding HDTV, then, of course, they will want to make the change. Their ability to do so depends very much on the availability and cost of the necessary equipment—cameras, VCR's, special effects, and other studio equipment, transmitters, etc. Virtually all this equipment must be newly purchased. Of course, the move to HDTV can be accomplished in stages, such as first simply passing through signals received from the network, then using taped or filmed productions, and finally, originating entire programs. This process will be quite expensive and will not be accomplished overnight.

During the transition period, the NTSC equipment must be kept running, as the market for HDTV broadcasting will grow slowly and simulcasting has been mandated by the FCC. Thus broadcasters face extra expenses for a long time to come. One problem they probably will not face is a shortage of program material. Virtually everything produced on film for NTSC is good enough for HDTV. This takes care of much of prime-time programming. Sports programs are another sure bet, as the wide screen and higher definition will add perceptibly to the visual effect. Of course, outside broadcasting equipment is needed for this function. Many current daytime programs really do not need HDTV and may well be aired in standard definition for many years to come, perhaps by using compression technology to fit several programs into one 6-MHz channel.

b) Equipment manufacturers: The Japanese companies that designed studio equipment to go with the NHK system

[8] If one were perversely designing an analog video system to achieve maximum interference, one would make all the transmitting systems scan in synchronism, like NTSC and PAL.

[9] Evidently, at the time that channel allocations were originally made there was not enough pressure on spectrum so as to mandate colocation of all transmitters within each city. With the reallocation opportunity provided by the shift to HDTV, this matter can be revisited.

are no doubt looking forward with great anticipation to the time when HDTV becomes a commercial reality so that they can begin to recoup their already very substantial investment. To some extent, the European manufacturers who did the same for HD-MAC will also be happy to make equipment for any system. Modification of their designs to accommodate a different coding system will cost much less than has already been spent on the design of cameras, monitors, VCR's, etc.

The situation with respect to receiver manufacturing is somewhat different, as the initial investment is much larger and the profit margins are much smaller than for professional equipment. Of course, the receiver manufacturers are also looking forward to HDTV broadcasting as opening a new market to them. In all likelihood, they will have little trouble finding the money required to enter the field, but they will be a good deal more cautious about committing to large-scale production until the level of uncertainty is reduced. Here price is the main factor, along with programming, that will determine the speed of penetration and therefore the possibility of making profits. Many observers think that an initial price of $3000–4000 would not be excessive. Both monochrome and color sets cost about that at today's prices when they were first introduced. The real question is whether HDTV receivers of, say, 35-in size, can be sold at that price, without losing money, within a year or two of introduction.

In NTSC sets, the cost of signal processing is negligible compared to the cost of display, cabinet, etc. That will not be the case with HDTV, as the processing power required far exceeds that found in today's most powerful personal computers. While there are many who argue that complexity is no longer a cost issue, the chips required for a system based on MPEG are exceedingly complicated. Pentium chips, for example, cost about $500[10] and they have much too small a capacity for real-time MPEG decoding. If HDTV is very successful, the volume should eventually exceed that of PC's. This is very much a "chicken and egg" problem in which it is hard to predict just what will happen.

c) Program producers: Like professional equipment manufacturers, program producers will probably be adequately motivated to get into HDTV as they see the market developing. Naturally, they will be influenced by cost considerations. In the case of 1125/60, which is already being used to some extent (although, except in Japan, the product must be converted to NTSC or PAL for broadcast), it is thought that concessional prices were offered by the equipment manufacturers in many cases.

d) Advertisers: Advertisers will certainly use any medium that brings them an audience, and will certainly not use any medium that does not. In the case of simulcasting, the total audience presumably will be only slightly more than would have been obtained with NTSC alone, so the total payment will only be marginally higher than for NTSC. It is conceivable that it will be found that certain kinds of advertising are more effective in high definition. In that case, advertisers will be more interested. In any event, it appears quite doubtful that advertising receipts can be counted on to pay for the transition to HDTV. When color was added to NTSC, RCA supported the new format to the extent of about $3 billion at today's prices. Who will provide the required investment this time is not clear.

e) Viewers: As mentioned above, $3000 would be an acceptable price for a large HDTV receiver, judging by earlier introductions of new systems such as NTSC color. In estimating the speed of market penetration, it should be recalled that it took 10 years to reach 1% penetration in that case, which was similar to the proposed transition to HDTV, since the same programs were seen in both formats. On the other hand, the receiver market today is very different from that in the 1950's. At that time, there were many domestic manufacturers, and many of these were making good profits. Intense competition has taken much of the profit out of the industry and caused most domestic manufacturers to go out of business.[11] It is therefore conceivable that it will prove impossible to create a mass market with receivers that cost so much.

There is another factor, however, which goes beyond price, and that is the relative attractiveness of the new and old formats in themselves, regardless of programs, which will be the same. Our own audience tests at MIT clearly showed that the relative preference for HDTV over NTSC, when both were shown with the same programs at studio quality, was small [7]. It seems obvious that the perceived difference would be much smaller than that between monochrome and color. However, we also found, indirectly, that there was a large perceived difference between studio quality, as used in the tests, and average quality in the home.

The decision to use digital transmission, about which the author has some serious reservations, does have a benefit in this case. With digital transmission, it is not possible to receive pictures that are seriously degraded by channel impairments.[12] *With NTSC, badly degraded pictures in the home are the norm.* Provided that adequate coverage and reliability are achieved with the all-digital system in the presence of the usual analog-channel impairments, and provided that compression itself does not produce serious impairments for a significant proportion of subjects, for the first time viewers will be seeing studio-quality images in the home. This is likely to be perceived as a substantially larger benefit than the higher definition. While it is a truism that viewers care much more about program content than about technical image quality, in this case they will see a

[10] On August 1, 1994, Intel reduced the price of 66-MHz Pentium chips from $750 to $525, in 1000 lots. Of course, TV decoders are unlikely to use completely programmable decoders in the forseeable future. This example is given only to show that very complex chips are not so likely to be cheap even in very large quantities.

[11] The only large American owned consumer-electronics company at present is Zenith, and that company does all of its manufacturing in Mexico. The largest manufacturers in the US are North American Philips and Thomson. The latter, owned by the French government, bought the consumer-electronics divisions of GE and RCA.

[12] Whether or not this is a benefit depends on how the overall system is designed. Extended coverage would be highly desirable even if there were some reduction in picture quality.

side-by-side difference in the store that may turn out to be important.

There are some who think that the 16:9 aspect ratio will be an important aspect of the appeal of digital TV. Of course, wide aspect ratio is also possible in analog systems, such as PAL Plus. There seems to be no good evidence that the wide screen is very important by itself. My personal opinion, which is shared by many in the creative community, is that the best aspect ratio is the one that was used to make the original production; e.g., portraits should be done in "portrait mode" and landscapes should be rendered in "landscape mode." In the focus groups used in the MIT audience-testing program, no evidence at all emerged that demonstrated that the wide screen, by itself, was a very important feature. The single parameter of the display that overshadowed all others, including sharpness, was image size.

4) An Acceptable Transition Scenario: In 1988, Zenith proposed a noncompatible HDTV transmission system that would use the taboo channels at low power, together with simultaneous transmission of the same programs on NTSC in current channels. Primarily on the basis of this proposal, the FCC decided to use simulcasting rather than a compatible signal format to serve existing receivers for a certain period. Broadcasters, who previously had been nearly unanimous in preferring a backward-compatible HDTV system, reluctantly went along. Ironically, Zenith's estimate of the adequate power level of the new stations was very far below what was later shown to be necessary. In addition, the source-coding method proposed at that time did not produce sufficiently good picture quality and was later abandoned. Nevertheless, the FCC stayed with its simulcasting decision, and eventually systems were developed that come close to meeting its requirements.

In one way, simulcasting solves the "chicken and egg" problem of noncompatible systems, in that the existing audience sees all the new programs, although not in HDTV. On the other hand, it removes much of the incentive to buy new receivers, since the old receiver permits viewing the new programs, just as if a receiver-compatible system had been used. It remains to be seen whether improvement in technical picture quality, by itself, will motivate consumers sufficiently to buy what are likely to be rather expensive new receivers. The alternative—attracting viewers to the new service by providing very desirable programs that cannot be seen any other way—was apparently rejected by everyone concerned as much too risky. My own opinion was that this course might have proven successful if a smaller and less price-conscious market, such as hotel television, had been tried rather than going immediately for the mass market.

In any event, the general idea of using simulcasting during the transition period is certainly feasible. That was the approach used in France and the UK when PAL was introduced in 1967. Old receivers were served for about 20 years, although not with all of the same programs made available on the new service. No one is immediately disadvantaged by simulcasting, but it does leave unanswered the question of how rapidly the public will make the shift to new receivers. If the FCC can stick to its intention of shutting down NTSC after 15 years, then, as that time approached, we would expect more sales of HDTV sets. One can expect the marketing of set-top converters from HDTV to NTSC to thrive, especially as, at least for some time, NTSC receivers will continue to be used with videotapes. Not only is 15 years a long time to wait for a market to develop, there remains some doubt whether Congress would allow NTSC ever to be abandoned if the public were strongly opposed.

A complete transition would mean discarding all NTSC equipment and making obsolete all existing receivers. An absolute necessity for this to be acceptable would be the availability of small inexpensive "HDTV" receivers, some portable, to serve the same functions that such receivers now serve. We do not want or need a theatrical experience while watching the morning news during breakfast, nor do the children need it for much of what they are now watching. We certainly do not want to pay very high prices for small receivers.

The main problem in making inexpensive sets to receive the HDTV signal is that, with existing American proposals, full decoding to baseband is required. The high-resolution image thus produced must then be processed to get the lower-resolution signal for the cheaper display. The need for a full decoder may well increase the cost of each set by several hundred dollars, and the selling price by even more. It would be better to have a coding system in which complete decoding were not required in low-performance sets. Even better would be a system with at least three levels of quality, with the cost of the decoder ranging from very low for the cheapest and smallest sets to substantially more for the full-quality receivers. This may well be feasible, but it is not part of the Grand Alliance proposal.

B. Regulatory Issues

Many aspects of TV system design cannot be settled by comparative testing; they must be decided on the basis of our preferences and the exigencies of the spectrum allocation problem. For example, coverage can be measured, but the aspect ratio must be decided upon on the basis of our preferences. The ability to function in the presence of a given degree of multipath can be tested, but whether we should deliver the same picture quality to everyone regardless of the distance from the transmitter is a policy issue. The amount of spectrum to be allocated to TV and the amount of service to be provided are basically political decisions.

1) What Kind of a TV System Do We Want? After about a half century of experience with television in the US, we have a good idea of its potential benefits and possibilities. Now that the time has arrived to have a new system, we have a rare opportunity to shape the medium in accordance with our collective views. Decisions on the overall nature of the service cannot be left entirely to the marketplace, since an enormous investment must be made before the public will get a chance to make its reaction known. Even in the current trend toward deregulation, no one has seriously suggested that transmission standards be left to the individual broadcasters or that spectrum assignments should no longer be made by the FCC. By setting standards and other ground rules, the Commission creates the environment in which the corporate entities that will provide service will function.

One good example of this kind of decision making is the support that the FCC gives to terrestrial broadcasters. Terrestrial broadcasting has immense support in Congress because it is the most used medium through which office holders get their message to voters. Many FCC regulations, such as the division of profits from reruns, appear to have been made with the primary purpose of keeping this industry alive.[13] Another example of regulation in the public interest, this time by act of Congress, was the All-Channel Receiver Act, which required all TV sets sold in the US to have UHF capability. This was a very successful example of government regulation of the free market that was to everyone's eventual benefit. Without it, many receivers would have been VHF-only, and the UHF spectrum would have proved impractical for TV.

2) The Need for High Spectrum Efficiency: NTSC has a very low spectrum efficiency. However, this is not due to stupidity on the part of its system designers. In 1941, when the standard originated, spectrum was not in short supply and cheap receivers had to have limited processing power. Neither of these conditions holds today. The electromagnetic spectrum is now a strictly limited natural resource. While the available spectrum is steadily being expanded at the upper end by advances in technology, TV occupies a large block of the more easily used UHF and VHF bands. In addition, it is now more practical to put a substantial amount of processing power into consumer products.

With the growth of mobile applications, pressure on the FCC to release unused UHF spectrum mounted. It was the fear of broadcasters that they might need more spectrum to compete with HDTV provided by alternative media that led to the current FCC inquiry that is working on HDTV standards. This has proved to be a very fruitful Inquiry, as it is leading to methods that are much more spectrum-efficient than NTSC. If the FCC's plan to turn off NTSC 15 years after HDTV broadcasting starts is actually carried out, we shall have at least the same amount of service as now within a considerably smaller spectrum allocation.

a) The role of source coding: It is obvious that if less bandwidth can be used for video of a given quality, or if quality can be improved without expanding bandwidth, the spectrum efficiency goes up. Until 1990 and the GI proposal, most executives in the TV industry thought that the first idea was impossible but the second might be accomplished. Of course, if one is true, the other must also be true, since these two statements are different ways of describing the same phenomenon, which is an increase in spectrum efficiency.

The method that has given the highest compression so far with manageable complexity, and is therefore used in all modern video coding systems, is the application of the discrete cosine transform (DCT) to the motion-compensated prediction error. Since provision must be made for scene changes and station switching, it is necessary to transmit some nondifferential information as well, either continuously (as the "leak" in DPCM) or from time to time. The net result is that the GA system can deal with no more than about three independent frames/s. For all its faults, uncoded NTSC can transmit 30 entirely independent frames each second, and each frame can comprise an arbitrary assemblage of sample values. The savings due to coding are dependent on successive frames being highly correlated and on each frame having high spatial autocorrelation (the efficiency of the DCT itself depends on the latter). While both of these situations are nearly always as stated, sometimes this will not be the case, and some new kinds of degradation will be evident [8].

b) The role of channel coding: One goal of channel coding is to fit as many programs as possible in each locality within the overall spectrum allocation for the service. This capability, although frequently ignored, is just as important as the compression achieved by source coding, which is universally recognized. In the US, at present, we can use about 20 channels in each locality out of 67 that are allocated, while in Britain the ratio is 4:44. Modern methods, as discussed below, may raise this ratio to 1:1. This would be just as important as reducing the bandwidth of a single program from 6 to 1.76 MHz!

The limitation of 20 out of 67; i.e., the existence of 47 "taboo" channels in each area, is due to a number of factors. The most fundamental, and hardest to deal with, is cochannel interference from another station on the same channel in an adjacent area. Given the carrier-to-interference ratio required for proper operation, the effective radiated power (ERP) of the transmitter, and the capability of a certain receiving antenna[14] and receiver, it is possible to calculate the minimum separation of stations, which is 160 mi for NTSC. This must be reduced to about 100 mi for HDTV in order to permit giving a second channel to each current broadcaster in accordance with the FCC's intended transition scenario. Clearly, HDTV must have much better interference performance than NTSC.

The second most important taboo is that adjacent channels cannot be used in the same cities, as discussed in Section II-A-2. The remaining taboos are predicated on poor receiver performance and are outdated. They need not apply to a new TV system.

[13] A topical example of government support for terrestrial broadcasting was the decision by the US Supreme Court on June 27, 1994, in which the economic viability of the broadcast industry was accepted as a legal basis for the reinstatement of the rule requiring cable companies to carry the local over-the-air programs. See L. Greenhouse, "Justices Back Cable Regulation," *NY Times*, June 28, 1994, p. D1.

[14] The antenna assumption is one of the "planning factors" established by the FCC to make it possible to calculate coverage area before a station goes on the air. The use of a better or worse antenna would make reception better or worse, but would not affect the calculation, which, to be useful, must be on a standardized basis.

3) Must All Receivers Have the Same Picture and Sound Quality? Television programs can be enjoyed over a wide range of image quality as long as the sound is free of serious distortion. At present, there is a wide variation of image quality from receiver to receiver. This is caused partly by differences in the size and quality of receivers and is also due to great variations of the amplitude and quality of received signals. The latter is affected by the kind of antenna used as well as by local conditions of signal strength, interference, and ghosts. These facts are widely recognized by the public as well as by TV professionals, although not often verbalized. No one, including the FCC, expects equally good pictures on all receivers; there is no FCC regulation of receiver image quality. On the contrary, should the FCC attempt to specify minimum receiver performance, there surely would be a storm of protest both from manufacturers and from free marketeers.

a) Receiver price versus performance: Typical households have two or three receivers. The best and largest is usually in the living room, while the others are in secondary locations such as the kitchen, children's rooms, etc. The latter, if bought for the purpose, are usually smaller and cheaper. While consumers certainly would not object to having maximum quality on all receivers, they have come to expect, as they do with most other products, that the cheaper sets will have lower performance. What would trouble consumers a good bit more would be the nonavailability of low-cost sets for these less critical uses.

In NTSC, it is possible for manufacturers to provide this range of price and performance because the main cost is the cabinet and display, compared to which the cost of the circuitry is almost negligible. This is not likely to be true with HDTV. Even in the largest and most expensive sets, signal processing will be an important part of the cost. If a complete decoder is required in all receivers, it will be the main cost in small sets. As long as this condition holds, it will not be possible to make inexpensive sets for today's less-critical applications.

This problem would be much less severe if simulcasting of NTSC were to remain in place indefinitely. However, the FCC's plan to take back a large proportion of the spectrum now allocated to TV requires abandonment of NTSC at some point. The lack of cheap receivers that can deal directly with the HDTV signal (or the lack of cheap set-top converters, which depend on the same technology) may prove an insurmountable obstacle to ever shutting NTSC down.

b) Portable and mobile receivers: While mobile receivers are not a big factor in the US, a very large proportion of sets in homes are portable in the sense that they may be moved from place to place and generally use on-set antennas—"rabbit ears." Well over half of the receivers in the US have antennas rather than being connected to cable or to satellite ground stations. This is a remarkable situation, since nearly two thirds of TV homes in the US [illegible].[15] What makes these ratios important is that the coverage performance of proposed HDTV systems is predicated on the use of a properly installed receiving antenna with 10 dB gain and 14 dB front-to-back ratio. One knowledgeable critic has even stated that, beyond 35 mi from the transmitting antenna, reliable reception will require a low-noise amplifier mounted on the antenna mast [9].

Under these conditions, it is clear that the abandonment of NTSC simulcasting will create a very difficult problem. Reception with rabbit ears will become unreliable, and coverage will be drastically reduced for receivers that do not have the assumed high-performance antenna. This will make it very difficult to maintain coverage and to provide low-cost receivers thus creating another obstacle to the FCC transition scenario.

4) Interoperability: Although there had been little talk of interoperability—the easy interchange of video data between systems of different performance, different applications, different industries, and different vintages—before it was raised in a very forceful way by computer interests [10], the frequent need for transcoding makes interoperability of great importance within the TV industry itself. The FCC eventually recognized this need by making interoperability a subject to be discussed in the Inquiry.

a) The need within the TV industry: Considering the large number of standards now in use and the still-unsolved problem of converting between NTSC and PAL,[16] one would have thought that it would not need FCC oversight to guarantee that transcoding would be taken into account during the design of a new system. Yet this was not the case. For example, the NHK system, which was the first format proposed for use as an international exchange standard, has scan rates that make it difficult to transcode either to PAL or NTSC.

The discussion in Section II-B-3 about the need for receivers with different price and performance illustrates that interoperability is not just a burden placed on the TV industry for the benefit of the computer industry, as is often stated. The ability to make simple receivers that can deal with a complex signal, even if their image quality is not as good as that of expensive receivers, is the key ability that is needed. It is so fundamental to system design that it cannot be added at a later date.

b) Nondisruptive improvement over time: Even before the computer industry was calling for an HDTV system that could easily be handled by workstations, the FCC itself was calling for "nondisruptive improvement over time." Learning from the NTSC experience, the Commission has

[15] These numbers are estimated from data provided by the Cable Advertising Bureau and Paul Kagan Assoc. Data from NCTA and Nielsen was also consulted.

[16] In spite of long effort, today's best transcoders are far from perfect, as was clearly demonstrated during the 1992 Summer Olympics. This event was shot in PAL and converted to NTSC for airing in the US. Defective rendition of rapid motion, such as disappearing volleyballs, was obvious, even though it detracted little from the popularity of the broadcasts. The reason transcoding is so hard is probably the prevalence of a great deal of temporal aliasing in all current TV systems.

made plain that any new system ought to be able to be upgraded without making earlier receivers obsolete. NTSC has very little room for progress in this way. The main change made since color was added in 1953 was stereo audio.[17] Any improvement in picture quality since 1941 is due to better cameras and picture tubes, and not to any change in system standards.

It does not take much reflection to show that, to improve the quality of a system after installation, it is necessary to send additional data that only new receivers would use. This data either must be hidden within the existing signal in such as way as not to degrade image quality on existing receivers or must be transmitted in a separate channel. In either case, many defects of the original system will remain in the enhanced system, even in new receivers.[18]

c) Across applications and industries: Interoperability became a public issue when it became apparent to the computer industry that the ability to display good-quality video on computer screens was very important to the future of the industry. With the still-declining cost of processing power, revenues can be kept up only by increasing the amount of computation. Nothing is so computation-intensive as high-resolution moving images. Even today's computers have a video screen, and many of the multimedia applications coming into use depend very heavily on video. It seems quite natural, therefore to display broadcast video on computers and to use computers to generate video sequences.[19]

Another industry that is affected is electronic imaging. Although no one thinks that film is going to disappear in the near future, it has become quite feasible to handle high-quality imagery in electronic form for virtually any application. Amateur photography is a good example. While equipment of full photographic quality is still too expensive for most users, properly handled images having a real resolution of 500–1000 lines are acceptable in many cases. If HDTV frames could be used as snapshots, an entire industry might be created. Similar possibilities exist in medical care, education, and publishing. The minimum demand of these non-TV industries is progressive scan and "square pixels." (equal horizontal and vertical resolution) What the TV industry is so far willing to give is all-digital transmission plus a self-description of each transmission by means of embedded headers and descriptors.[20]

[17]Since the addition of color substantially reduced the luminance resolution of receivers, existing or to be manufactured, and added cross color and cross luminance to the jargon, one would have to say that the 1953 changes, while praiseworthy, were not entirely "compatible."

[18]The extreme vulnerability of NTSC to interference and the associated poor spectrum efficiency as well as all the disadvantages of interlace, are related to its system design and cannot be cured by improved receivers. Ghost cancellers might well improve the performance of new NTSC receivers. The system described in Section III-D is specifically designed to permit upgrading over time.

[19]Computers are already widely used to create and edit video in the NTSC format. Unwieldy as it is, it has nevertheless proved quite feasible to design the hardware and software needed for this application.

[20]The TV industry is not a monolith on this or any other question. For example, ABC and FOX, two of the four TV networks, favor progressive scan.

C. The Transmission of Media and Their Characteristics

In the US at present, video signals are transmitted to receivers by means of terrestrial (over-the-air) transmission, by cable, by VCR, and by satellite. Each of these media has different physical characteristics that must be taken into account in order to get the best results. The last is by far the least important in the US, since it is confined to a few million users who tune in directly on the programs being sent to TV stations and to cable head ends. However, this year a satellite has been launched and two operators, DirecTV and USSB, are providing service. Initial acceptance has been good, so the situation may change.

1) Terrestrial Transmission: Terrestrial transmission is the most popular medium in terms of receivers served. It is free in the US and widely used for political purposes, giving it immense support from the public and in Congress. Technically, it is the worst medium, suffering from noise, ghosts, interference, and frequency distortion. A unique characteristic is the very wide variation in signal strength from receiver to receiver. Coupled with the differences in receiver noise performance and antenna characteristics, a very wide variation in CNR is encountered, corresponding to more than a 5:1 range of channel capacity. The NTSC signal design is such, however, that good synchronization and good audio quality are maintained under virtually all conditions in which the image is even marginally viewable. Very simple antennas can be used except at the boundary of the service area. In the absence of interference, with a good receiving antenna, and with a line of sight to the transmitting antennas, programs can be viewed some 200 mi from the transmitter site.

Twelve VHF and 55 UHF channels are allocated for TV, with a maximum of seven VHF and about 12 UHF stations actually licensed in each city.[21] Adjacent channels are not used in any one locality and stations on the same channel must be at least 160 mi apart. Broadcasters greatly prefer VHF assignments, since better coverage is obtained with lower transmitter power. In the absence of cochannel interference, and using the maximum permitted ERP, coverage is noise-limited somewhat beyond the radio horizon—52 mi for an antenna 1350 ft above the ground (HAAT). In certain areas of the country, HAAT's of as much as 2000 feet may be used. This has a radio horizon of 63 mi, but a noise-limited range of 80 (channel 2) to 67 (channel 69) mi. Actually, few stations have maximum-height antennas.[22]

2) Cable: Cable service is available to about 96% of the 95 million TV homes in the US and about 65% actually subscribe. Although cable provides a much larger number of programs than terrestrial broadcasting, most cable viewing is of programs that originate with the networks. In principle, all of the technical problems mentioned in connection with over-the-air transmission ought to be absent on cable, but they are not.

[21]On average, each television household in the US has 13.3 free stations available to it (Nielsen).

[22]Information on antenna height from Dr. T. J. Vaughn of Micro Communications Inc.

At present, cable uses trunk and branch distribution, with amplifiers along the trunks as needed. Some nonlinear distortion is introduced in this way. Coaxial cable is almost always used into the residence, but fiber is steadily replacing cable on the trunks. Cable is not completely impervious to leakage either in or out, so the same kind of natural and man-made noise is encountered as in terrestrial broadcasting, although to a lesser degree. Passive lossy signal splitters are used in many locations, with unused taps generally unterminated. This creates a kind of endemic multipath that behaves much like a low-pass filter.[23]

Signal strength from receiver to receiver is more uniform than over the air, but still varies because of the use of signal splitters. All channels have signals of about equal amplitude, so that there is no adjacent-channel taboo as in terrestrial. Cable companies try to ensure 38–40 dB CNR at the receiver terminals, but do not always succeed. If they did, the noise would be marginally visible but not annoying. In spite of all this, "cable quality" is generally superior to average quality with rabbit ears. In many locations, however, a good antenna produces better quality than provided by cable. Informed opinion is that viewers usually subscribe to cable because of a wider choice of programs, and not for higher image quality.

3) Video Recorders: For every two receivers in American homes, there is one VCR.[24] Although originally used mainly for time-shifting, the vast majority are now used for playing rented movies.[25] There are also about 22 million camcorders. Thus, tape viewing accounts for a significant portion of TV use. Any new system *must* have affordable and reliable VCR's to be acceptable.

Getting two hours of NTSC signal onto a small spool of tape was a remarkable technological achievement that required some compromises with signal quality. Sometimes, "VHS quality" is used as a measure to indicate something considerably below that of NTSC. Certainly, the resolution and SNR of the VHS format is lower that of studio-quality NTSC. However, NTSC as typically viewed in the home is also quite inferior to NTSC in the studio. My own opinion is that with a good tape and a VCR in good condition, one gets better pictures, on average, from tape than from broadcasts.

4) Satellite Broadcasting: In principle and in practice, the satellite channel is substantially superior to all other existing means of transmitting video to the home. A line-of-sight path is always used, along with directional antennas. There is very little multipath and little adjacent-channel interference. Cochannel interference would be much like that of terrestrial broadcasting from a single centralized antenna. Most current transmission, which was never intended for broadcasting, is analog FM using an RF bandwidth of 36 or 54 MHz. This gives a favorable "triangular" noise spectrum. Some digital transmission is also used with a very conservative data rate of only 45 Mb/s. The system noise budget is arranged so that even under extreme weather condition such as heavy rainstorms, the received signal is well above the threshold, and reception is studio quality.

For DBS to the home, a bandwidth of 24 MHz will be used. For the less demanding requirements of home reception, it will most likely be found that a gross data rate of some 60 Mb/s per channel can be used as compared with 20–25 Mb/s for terrestrial broadcasting. This will permit transmission of two HDTV signals or 8 standard-resolution signals, with far higher reliability than is likely to be experienced with terrestrial transmission.

III. Some Possible Solutions

A. Source and Channel Coding

Shannon's work can be interpreted to mean that source and channel coding ought to be independent. In this approach, the source coder removes all statistical redundancy, producing a signal that looks like random noise; the channel coder adds redundancy in just the right way so as to permit near-perfect error correction. Each coded bit is then essential to reconstruction. However, such a scheme is impossible to implement exactly, since all redundancy cannot be removed. If it were, a single error would make further decoding and resynchronization impossible. The closer we get to such an "ideal" system, the more fragile the signal, the longer the coding and decoding delays, and the more difficult the synchronization.

In the best current systems, the data transmitted is very far from being equally important. In addition, the concept applies only to point-to-point systems in which the receiver CNR is well defined. It does not apply to broadcasting, in which very large differences in CNR are found from receiver to receiver.[26] Thus terrestrial broadcasting requires a rethinking of the coding problem if optimum use is to be made of the limited spectrum that is available.

There are two approaches that can be taken. Using high-power centralized transmitters as at present, one solution involves self-optimization at each receiver according to the amount of data that can be recovered. The latter should be as close as possible to the Shannon capacity at that receiver. Necessarily, everyone does not get images of equal quality. The second solution involves making the signal strength, and therefore the channel capacity, as nearly uniform as possible across the population of receivers. This can be done by using a cellular network of low-power transmitters, all emitting the same program. If the transmitters in the cellular network all operate on the same frequency, the arrangement is called a single-frequency network (SFN). The receiving area can be delineated almost arbitrarily by the placement of the transmitters, and contiguous areas can use the same channel for different programs. This

[23] In the US, it is not unusual to find ghosts on cable similar to those encountered in over-the-air reception. In most cases, these ghosts were present in the signal when received at the cable head end.

[24] Data from Zenith Electronics Corporation.

[25] [illegible]

[26] The broadcasting problem, unfortunately, has attracted very little

method achieves the highest possible spectrum efficiency cochannel interference disappears as a design issue. Only as many channels need be allocated to TV service as the number of independent programs that are to be available in each locality.

1) Multiresolution by Combined Source and Channel Coding: In analog systems, image quality necessarily deteriorates steadily with falling signal quality, primarily through lower SNR. The resulting soft threshold can be thought of as a rough kind of self-optimization (The sound quality remains good at a signal level that produces barely watchable images, and that is probably a good choice to make in new systems). To achieve the very high compression ratio needed to transmit HDTV in a 6-MHz channel, at least some digital data must be transmitted. In digital transmission, there are no known methods of getting a soft threshold, i.e., of recovering a continuously higher digital data rate from a continuously rising CNR.[27] Thus recovery must be a stepwise affair. This means that the source coder must organize its output into a number of data streams in which the quality increases with the number of streams recovered. The channel coder must package these data streams in the transmitted signal in such a way that the number of streams recovered increases in a stepwise fashion with receiver performance and with the signal strength at the receiver terminals. Finally, the receiver must make the best possible picture from the recovered data at each level of CNR.

Resolution and SNR are the two image-quality factors that depend on the amount of data recovered. There is no consensus as to which should be varied the most from level to level; MPEG2 provides both possibilities [12]. A small amount of white or high-frequency noise is relatively harmless, but an amount and character of noise much different from what is now seen when reception is deemed acceptable is probably unwise. On the other hand, there is clearly a very large tolerance for resolution differences, as today's situation makes obvious. This is not only true for small receivers, which look sharp even when the resolution in absolute terms (number of samples per picture dimension) is quite low. It is also true for large displays. Their resolution in absolute terms is quite low, but they are nevertheless preferred. In audience tests at MIT, image size was by far the most important factor in viewer preference [13]. Viewing angle, which is of great importance in subjective assessment of TV displays, cannot be controlled by the system designer.

These observations provide enough direction for designing a system using several levels of quality. We shall designate such systems as using multiresolution (MR) coding as distinct from single-resolution (SR) coding, even though both resolution and SNR may vary from level to level.

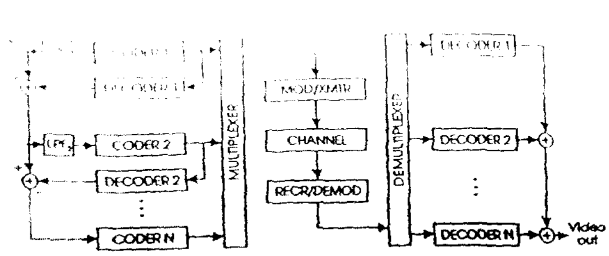

Fig. 1. *Pyramid Coding.* This is the basic arrangement of a multiresolution system that provides good picture quality at every level of performance. A low-pass filter (2- or 3-d) selects information that is to be included at the lowest-quality level. This is coded and decoded and then subtracted from the original video. A second low-pass filter provides information for the next (enhancement) level, which is also coded, decoded and subtracted from the remaining input video, etc. (Subtracting decoded data at each level ensures that any coding distortion is available to the next higher level for possible correction.) The coded data streams from all the levels are multiplexed, modulated, and transmitted. The receiver combines the decoded lowest level with whichever enhancement levels are recovered to produce the best picture that can be made from the available data.

a) Multiresolution source coding: There is a considerable literature on MR systems, as they are useful in a number of applications, including browsing through image data bases.[28] An early paper coined the term "pyramid coding" for schemes in which a basic image could be upgraded by addition of more information, as shown in Fig. 1 [15]. The general idea was used in a number of proposed receiver-compatible HDTV systems for the US in which enhancement data, either hidden within the main signal or transmitted in a second channel, would be added to a standard NTSC signal [16].

A significant aspect of pyramid coding is that, to be useful, all the pictures in the hierarchy must be free of obvious defects such as ringing (Gibbs phenomenon) due to sharp-cutting filters. To avoid this problem, the filters that separate the several data streams must have a smooth and not-too-rapid cutoff. As a result, the same frequency component may be represented in more than one stream. With existing coding technology, this results in a penalty in the quality/compression tradeoff as compared with systems that code the entire image spectrum in one stream. In general, pyramid systems require a somewhat higher data rate at their highest level to achieve the same quality as that of SR systems. This is offset by the ability of MR systems to provide good pictures, albeit of lower resolution, at lower data rates which permit greater coverage. MR systems can also provide higher quality than SR systems when it is possible to deliver more data to the decoder.

b) Multiresolution channel coding: For digital transmission, it is sometimes suggested that unequal error protection can be used to achieve multiresolution [17]. However, the numbers do not work out very well. The amount of error protection required at low CNR is very large and leaves little room for the real data. Another proposal is to subdivide the channel by frequency or time, using constella-

[27] In [38], the authors describe a spread-spectrum method that produces a quasi-continuous threshold for the channel coder. It is not clear whether adding transform coefficients in a quasi-continuous manner will give good picture quality at all levels.

[28] This was sometimes called "progressive transmission," which must be carefully distinguished from progressive scanning [14].

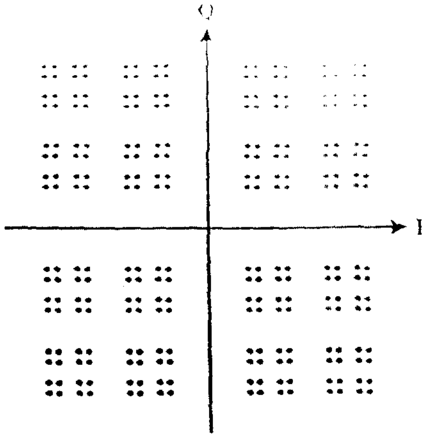

Fig. 2. *Nonuniform Constellation.* **This constellation has four levels of performance with CNR thresholds approximately 6 dB apart. It is intended to be used with a multiresolution source-coding method that produces four streams of data.**

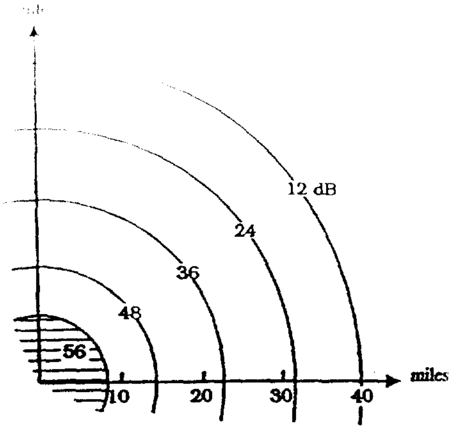

Fig. 3. *Variation of CNR With Range.* **The inverse-square law does not govern typical TV antenna performance. This is because of its height and the vertical profile of its beam, as well as high attenuation at the edge of the service area. Grazing incidence in this area causes the field strength to diminish very rapidly with distance. Finally, the FCC planning factors, which rise with distance, effectively reduce the field strength, producing the result shown. The most notable features are the near-uniform field strength in the inner 8 mi and the uniform decrease in signal (in dB) with distance. Note that the channel capacity, which is proportional to CNR in dB, is more than four times as high downtown as at the threshold of service. (Data from Dr. O. Bendov.)**

tions of different density (different numbers of bits/cycle) in the various subchannels. This is also inefficient, since at the threshold CNR for a dense constellation (i.e., finely quantized), subchannels with less dense constellations (coarsely quantized) are very inefficient. At the present time, the best known method is to use a multilevel modulation scheme such as the nonuniform constellation as in Fig. 2.

As is the case with MR source coding, MR channel coding is also somewhat less efficient than SR coding at the design threshold of the latter. However, the MR system becomes more efficient than the SR system at higher CNR. In addition, the former can deliver pictures, albeit of lower quality than that of the latter, at substantially lower CNR, thus extending the coverage area.

c) Overall performance of MR systems: The variation of receiver CNR with range for a typical current-day UHF transmitting antenna is shown in Fig. 3.[29] Note that the channel capacity, which is proportional to the CNR, decreases by a factor of more than four from the central to the outlying area. Obviously, sending the same data rate to all receivers wastes a great deal of capacity in just those close-in areas where spectrum is in shortest supply.

In Fig. 4, a comparison is made between the performance of SR and MR systems, in which the design threshold of the former is 16 dB. In such an SR system, an HDTV image of uniform quality is delivered everywhere the CNR is at least 16 dB, and no picture at all is delivered beyond that. In the MR system shown, a low-resolution image is delivered from 6 to 16 dB, a medium resolution image at 16–26 dB, an HDTV image similar to that of the SR system at 26–36 dB, and a better-than-HDTV image for CNR's in excess of 36 dB. In qualitative terms, the MR scheme extends the service area considerably beyond that of the SR system and delivers superior pictures for CNR's higher than 36 dB. The price paid is a reduction in quality for CNR's between 16 and 26 dB. While these numbers are not associated with any particular system, they are believed to be typical.

2) Single-Frequency Networks: Although the SFN concept is not new, it was recently brought to prominence by its proposed use in digital audio broadcasting in Europe [18]. It is also used in some radio applications [19]. The entire service area of a station can be covered with a cellular array of same-frequency low-power transmitters, or the array can be used in the outer region and a single medium-power transmitter, or even a satellite broadcast, can be used for the central region. The various transmitters may be fed by cable or in a different channel, or all transmitters may derive their signals from each other. The carriers may be identical or intentionally offset. Some successful field tests have been carried out, but no full-scale SFN has yet been implemented. There is considerable controversy over details of the expected performance [20].

Within the cellular array, the signals from a group of nearby transmitters appear as multipath at the receiver. The amount of multipath can be reduced, but not eliminated, by use of directional antennas [21], but it would be far preferable to use simple antennas, perhaps omnidirectional, in a large percentage of locations. Thus the multipath performance of the modulation and channel-coding system emerges as a principal concern. Multipath is a linear distortion, equivalent to the effect of a certain filter. Its two main effects are intersymbol interference (ISI) and a possible increase in noise level due to equalization of the multipath distortion.

[29] This diagram takes account of the "planning factors" used by the FCC in determining coverage. Among other things, these factors deal with the percentage of times and percentages of locations in which the given reception conditions are met or exceeded. In the central area, signal strength is nearly constant. This is due to the vertical profile of the transmitter's antenna beam and to the fact that the receiving antennas are much closer to the ground.

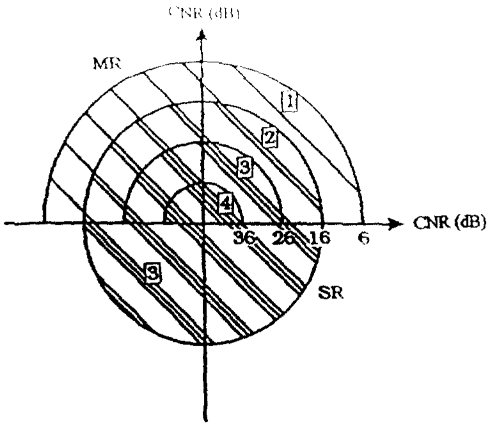

Fig. 4. *Comparison of Typical Single- and Multiple-Resolution Systems.* The thresholds are shown in circle and the quality levels in squares. The SR system provides the 3rd level of quality everywhere where the CNR is >16 dB. The MR system provides a larger service area (out to 6 dB) at lower quality (1st level) and higher-than-SR quality (4th level) where the CNR >36 dB. The price for this improved overall performance is lower quality (2nd level) between 16 and 26 dB. The two systems have the same quality from 26 to 36 dB. The numbers here do not represent any particular MR system; they are intended to show a typical relationship between the service rendered by an MR and a SR system using compression schemes of roughly equal effectiveness.

While the main advantage of SFN's is spectrum efficiency, there are other advantages as well. Service areas can be of irregular shape, and can include regions that are otherwise denied reception because of intervening obstructions. Except for a narrow region along the boundary of the service area, the transmission power can be raised enough so that CNR is no longer a factor in reception. Even so, the total emitted power is much less than that needed by a single centralized transmitter.[30] Note that the improvement in spectrum efficiency due to MR coding is less important in SFN's than in the conventional single-transmitter arrangement. However, the facilitation of the manufacture of receivers of a range of price and performance makes MR coding advantageous in all cases.

ISI due to multipath reception can be removed by equalization or by use of multicarrier reception as discussed below. The accuracy, complexity, and noise performance of these schemes are the main issues.

3) Multicarrier Modulation: The distorting effect (the ISI) produced by a given level of multipath depends not only on the total power and relative delay of the echoes but also on the ratio of the temporal spread of the echoes to the symbol length of the signal. In VHF and UHF terrestrial transmission, most echoes occur within about 20 μs of the main signal. This does not cause much trouble with AM or FM audio broadcasting, with a symbol length of about 25 μs, but it produces heavy impairment in television, with a symbol length of about 120 ns. Obviously, one way to reduce (but not eliminate) the distortion is to divide the signal into a large number of components, each of which has a much longer symbol length, and to transmit these components as narrowband modulated carriers within the original channel. The ISI can be eliminated completely by inserting after each symbol a guard interval during which a portion of the symbol waveform is replicated. This permits integrating each symbol over its symbol duration without unintentionally including energy from symbols just before or after the symbol being demodulated. The guard interval itself must be longer than the multipath spread. Since the guard interval reduces the efficiency of the transmission, it is advisable to make the symbol long as compared with the guard interval, with a correspondingly large number of carriers.

Frequency-division multiplex, as discussed above, has been improved by two developments—orthogonalization of the modulated carriers so that no bandwidth need be wasted by using guard bands, and implementation by means of the discrete Fourier transform [22]. The resulting system, including coding, is called coded orthogonal frequency-division multiplex (COFDM). It is already used in some modems for digital data transmission over telephone lines, and is being planned for use in digital audio broadcasting in Europe [23]. It is the subject of a companion paper in this journal [24].

Another important property of OFDM is that out-of-band radiation is much less than in single-carrier modulation (SCM). This is because orthogonality, as produced by the discrete Fourier transform (DFT), makes the spectrum of each modulated carrier have the shape $(\sin(\omega)/\omega)$ centered on the carrier frequency, with the zeroes placed at the locations of the neighboring carriers. With hundreds, or even thousands, of carriers, the spectrum thus decays extremely rapidly at the edge of the channel, even without filters.

The elimination of ISI by OFDM, although very valuable, is not a complete solution to the transmission problem, as we must still deal with the noise caused by equalization of the multipath channel. Originally, the claim was made that COFDM adds echo power constructively, so that the error rate actually goes down with more echoes. While it is true that, averaged over all receivers, the powers of signal and ghosts do add, this is not true at every individual receiver (The BER goes down in some cases and up in others). Depending on the precise character of the echoes, deep notches may be produced in the spectrum. The worst case is that of a single echo of 0 dB, which produces actual nulls. Data transmitted on carriers at frequencies where the signal strength is very small is obviously less reliable. This can be dealt with by interleaving and coding, but it is clear that, at some locations, transmission may be adversely affected. One remedy is the use of directional antennas at those locations. In most cases, these would not have to be very elaborate, as it is only necessary to reduce the offending ghost by 3–6 dB. Simple dipoles would suffice in many cases.

Wideband nulls can also be caused by radio-frequency phase cancellation. A solution in most such cases of this

[30] Single transmitters are remarkably inefficient in covering large areas on account of the very rapid decrease in signal strength with distance near the boundary of the service area. It takes an increase in power of [illegible] to 1.5 dB to increase the range by 1 mi.

kind is simply to move the antenna by a fraction of a wavelength. More elaborate installations could use space diversity reception.[31]

The tradeoff in complexity between receivers for SCM and for OFDM involves the time-domain equalizer used in the former versus the DFT required for the latter. In OFDM, a frequency-domain equalizer, which is far simpler than a time-domain equalizer, is most natural. On the other hand, OFDM requires the DFT operation, which is not needed in SCM.

4) Digital versus Hybrid Transmission: In the "ideal" system discussed in Section III-D, we use hybrid analog/digital transmission. This undoubtedly seems a quaint idea from the past to those who have joined the digital bandwagon. However, careful analysis of some specific aspects of coding systems shows that digital transmission does not have all the advantages claimed for it. It is true that some digital data must be transmitted in order to achieve the very high compression associated with motion-compensated transform coding. However, it is also true that higher channel-coding efficiency can be achieved with hybrid transmission. Finally, interoperability is not materially enhanced by all-digital transmission.

a) Source-coding efficiency: In motion-compensated transform coding, the amplitudes and identification of adaptively selected transform coefficients comprise the bulk of the data to be transmitted. In the GA system, this data is jointly coded for 2–3 million coefficients per second at about 4–6 b/coefficient. In fact, the nature of the large correlation between amplitude and identification (the spatial frequency of each selected coefficient) is such that not much would be lost by separately coding the two kinds of data. (This is discussed further in Section III-D-1.) If the statistical relationship among the coefficient amplitudes themselves is not utilized in the coding scheme, there is nothing to be gained by quantizing the amplitudes before transmission. That simply adds quantization noise. Analog transmission works well in this case. The data that must be transmitted per coefficient in a hybrid system is one analog sample plus less than one bit. All other aspects of MPEG coding can be used with hybrid analog/digital transmission, so that comparable compression ratios can be achieved.

b) Channel-coding efficiency: In Section II-A-2, we pointed out that, when analog information (such as the amplitude of transform coefficients) is sampled and quantized for digital transmission in an analog channel, the requirements for achieving a transmission rate close to the Shannon rate include very fine quantization combined with very effective error correction. Note that noise added to these coefficients produces no catastrophe in the reconstructed image; thus, they need not be entirely noise- and error-free. The requirement for near-perfect transmission in MPEG-like systems arises from joint coding of the amplitudes with the adaptive-selection data, for which errors produce serious image defects. On the other hand, analog transmission of the coefficient amplitudes can readily achieve the full Shannon capacity, and it can do this for a range of CNR, and not only the threshold CNR. For the peak-power-limited additive-white-noise analog channel, if the coefficients comprise a train of uncorrelated analog samples of uniform amplitude probability distribution, the mutual information (i.e., what the noisy output signal tells us about the noiseless input signal) is equal to the Shannon capacity of the analog channel in which they are transmitted. (For an RMS-power-limited channel, a Gaussian distribution is optimum.)

Since the coefficients to be coded represent differential data, i.e., prediction error, and must therefore be integrated to generate the desired output, it may be thought that analog transmission cannot be used because of the possibility of a catastrophic accumulation of noise in the decoder output. The coefficients in their analog form have precisely zero average value, as does the channel noise. The average is approached fast enough so that no catastrophe occurs, as we have demonstrated in our simulation. The "integrator" in this case can have zero response at zero frequency and still produce the desired output.

c) Interoperability: The difficulty of transcoding between two different video signals is primarily a function of their relative sampling grids. It makes little difference if the signals are in digital or analog form, since conversion from one form to the other is rather simple. If the signals are compressed, it is generally necessary to convert to uncoded form to do any transcoding at all.

The fact that the two systems have different spatial sampling frequencies does not present much of a problem since the sampling theorem provides the theoretical basis for moving from one grid to another. In practice, filters should be chosen with due regard for perceptual effects [25]. Different temporal sampling rates, however, always cause trouble. This is because temporal aliasing is nearly always present unless motion is less than one sample/frame. The aliasing greatly inhibits temporal filtering, which is prone to produce defects such as multiple images. With the amount of motion commonly encountered, a rate of even hundreds of frames/s is insufficient to allow the elimination of temporal aliasing without excessive blurring. Blurring of moving objects is counted as a defect to such an extent that electronic shutters are sometimes used although this makes the aliasing worse.

Good temporal interpolation can only be done if motion compensation is used. While this is quite complex, good results can be achieved. In Ph.D. dissertations by Martinez and Krause [26], essentially flawless transcoding was demonstrated with arbitrary ratios of frame rates.

Another factor in interoperability is the complexity of the relationship between the transmitted signal and the uncoded video signal that it represents. *High compression*

[31] A single echo causes the frequency response to undulate over the band with a frequency separation between peaks equal to the reciprocal of the relative delays. If the relative delay is comparable to the reciprocal of the radio-frequency (RF) bandwidth, a single cycle of the undulation is about as wide as the rf band. Assuming that the signals come from different directions, the null can then be moved a great deal by shifting the antenna on the order of one wavelength. In general, the antenna has to be moved on the order of the velocity of light (c) multiplied by the relative delay. The exact amount depends on the directions of the signal. For relative delays of more than λ/c, antenna diversity is not practical.

ratios necessarily involve complex coding algorithms. If it is necessary to decode an HDTV transmission completely in order to extract a low-resolution video signal for display in a small low-performance receiver, the receiver cannot be so low in cost. It is much better to use a pyramid coding scheme in which the simplest receivers deal only with the lowest level of the pyramid and can therefore use the simplest and least expensive decoder.

Interoperability is also affected by the channel coding scheme. Ideally, one would like a range of encoders of different quality (resolution) to be able to communicate with a range of decoders. In this way receivers of different price and performance could all accept the same transmitted signal, while the signals transmitted from a range of encoders of different resolution would all be acceptable by all decoders. One way in which this can be done is discussed in Section III-D.

B. Noise and Interference Control

Noise can usually be defeated by transmitting at higher power, although some limits are set by practical and economic considerations. However, the main limitation on transmitted power comes from the need not to interfere excessively with other stations. In the case of HDTV, the FCC's intended transition scenario calls for adding HDTV stations while current NTSC stations remain on the air. This must be done without materially reducing the latter's coverage, while at the same time attaining adequate coverage for the new transmissions. After NTSC is shut down, only HDTV stations will remain on the air, and they must have coverage similar to today's stations, but within a reduced overall spectrum allocation. It is clear that HDTV signals must be recoverable at lower CNR than now required for NTSC and that they must have better interference performance. To the extent that digital data is transmitted, error correction and concealment must be implemented in order to achieve appropriate image and sound quality. To the extent that analog information is transmitted, the recovered signals must have appropriate SNR.

For best noise performance in the additive white Gaussian noise channel, the spectrum of signals should be uniform.

1) Noise Performance for Digital Data: Within a given channel capacity as limited by bandwidth and CNR, errors caused by noise are correctable, in principle, by coding, as long as the Shannon rate is not exceeded. The closer the total transmission rate (signal data plus error-correction data) to the Shannon channel capacity, the higher the uncorrected (raw) error rate. To achieve net transmission rates that are a substantial fraction of the Shannon rate, the raw error rate must be quite high. A combination of outer Reed/Solomon plus inner trellis coding has proved to be an effective method with manageable complexity and coding delay [27]. A corrected bit-error rate (BER) of 5×10^6 is the generally accepted threshold of service, as error concealment is effective at that rate.

All digital modulation methods have sharper thresholds than analog schemes, and coded digital methods have extremely sharp thresholds. In analog systems, which have soft thresholds, coverage is usually calculated on the basis of a CNR that is exceeded in half the homes half of the time. There is as yet no generally agreed-upon values for these percentages for digital transmission, but it is clear that reception must be guaranteed much more than 50% of the time.

2) Noise Performance for Analog Data: In uncoded analog systems such as NTSC, the SNR of the recovered video signal is exactly equal to the CNR of the transmitted signal. In coded analog systems, such as FM or spread spectrum, it is possible to trade off bandwidth and SNR, although the tradeoff is generally not as effective as in digital modulation such as PCM. If the bandwidth of the data to be transmitted is less than that of the channel, an improvement in SNR can be achieved. For example, if 5 MHz is the usable channel bandwidth, 10^7 samples can be transmitted per second. If the number of samples to be transmitted is less than this, the SNR of the recovered signal can be higher than the channel CNR. With spread spectrum, if the different original signal samples require different SNR, then another improvement is possible by transmitting the more sensitive samples at relatively higher power without changing the statistical parameters of the signal in the channel [39].

3) Interference Performance: For a given relative power, analog signals interfere the least with each other when they appear to be random noise to each other.[32] This is easily accomplished with digital transmission, and is one of its major advantages, but rarely mentioned. One result is that the threshold carrier-to-noise ratio is about the same as the threshold carrier-to-interference ratio (CIR). Analog signals must be scrambled to accomplish the same end, and this is also readily accomplished with modern technology.

During the transition period to all-HDTV broadcasting, the interference between HDTV and NTSC is an important consideration. Interference is mutual; If A is less interfered with by B, it can be transmitted at lower power, thus interfering less with B. Of course, reducing power may reduce coverage where it is noise limited. It is much easier to plan the location and power levels of transmitters when no stations are already on the air in the band in question. When adding HDTV stations in the spectrum now allocated to NTSC, the problem is much more difficult. However, strong resistance to noise and interference is always helpful.

4) Synchronization and Accurate Carrier Recovery: Although not a factor in spectrum efficiency, synchronization of all clocks is a very important practical consideration. Accurate clock recovery is vital to minimizing the BER. The ability to synchronize rapidly and accurately in the presence of noise, multipath, and interference is essential to achieving proper coverage and is a great convenience when changing channels. One of the merits of NTSC is its ability to synchronize under very noisy conditions, a merit

[32] This is one of the most serious limitations of NTSC. Relative randomization of the scanning patterns would have greatly improved the interference performance. On the other hand, the known nonuniform spectrum of NTSC can be used to decrease its interference into fully randomized signals [28].

that is not surprising since more than [illegible] of the channel capacity is devoted to this purpose.

In principle, synchronization does not require the use of any channel capacity. If the system is well designed, statistical parameters of the signal, such as RMS value, autocorrelation function, etc., are well determined and can be used for this purpose. The use of synchronization signals not only uses some channel capacity, but inserts some periodicity into the signal, which increases its potential for interference with other signals. As a practical matter, and in view of the current state of the art, it appears than devoting a small amount of channel capacity to this function and accepting a slight increase in interference are defensible decisions. In the GA competition for the channel-coding scheme, the Zenith system, which does use pilot carriers, was able to synchronize at substantially lower CNR than the GI scheme, which did not. This was an important factor in choosing the former over the latter [31].

C. Multipath and Frequency Distortion Control

Multipath, which is a linear distortion, can be corrected by linear equalizing filters in the same manner as other sources of frequency distortion. Noise limits the performance of equalizers in two ways. If the uncorrected signal is noisy, calculation of the filter parameters must be done slowly enough so as to average out the noise. Even if the filter parameters are correct in terms of frequency response, a large increase in noise may result if there are near-nulls in the uncorrected spectrum. For SCM, errors are caused both by incompletely corrected frequency response, which leads to an imperfect "eye" pattern, or by noise, which also partially closes the eyes.

Echoes can be reduced in amplitude, but generally not completely removed, by use of highly directional receiving antennas. Almost whatever modulation and error-correction systems are used, it probably will always be necessary to use directional antennas at those locations that otherwise would have near nulls in the spectrum.

The situation is somewhat different in multicarrier modulation (MCM) because the data on carriers received at relatively low amplitude has a higher BER than data on carriers received at relatively high amplitude. The data in each transmitted block can be distributed across many carriers (preferably all of them) and the performance linked by a code. For example, the portion of the data with lower CNR can be weighted less heavily by the decoder [30].

There is very little data available on the effect of equalization on CNR in typical broadcasting situations. Recent tests at the Advanced Television Test Center using seven different combinations of echoes with a total power 7.5 dB below the direct signal have shown that the threshold CNR goes up, averaged over the seven echo sets, about 2.5 dB [31]. It should be kept in mind that much worse echoes are often encountered and that, therefore, a substantial reduction in coverage is likely if there are large echoes near the boundary of the service area.

1) Implementation of the Equalizer: Equalization can be carried out in the time domain or the frequency domain. In the time domain, an FIR filter somewhat longer than the temporal spread of the echoes is effective in most cases. The output is a linear combination of the signals at the various taps of the filter—typically 256 to 1024. The tap coefficients are obtained by various methods. Sometimes clock recovery is combined with coefficient calculation. Some methods use transmitted reference signals and some ("blind deconvolution") use the main received signal itself as reference [32].

In the frequency domain, equalization can be accomplished by dividing the channel output into a large number of narrow-band components and multiplying each by a single complex factor. This method is based on the assumption that the frequency response is constant across each narrow band, which is almost certainly justified when there are many hundreds of channels. The effect of such an equalization is exactly the same as that of a corresponding linear filter operating in the time domain. Note that in this form of equalization, a convenient pilot signal consists of an assemblage of sine waves or a swept-frequency signal, sometimes called a chirp. A convenient pilot signal for time-domain operation is one that determines the impulse response of the channel, such as a pulse.

Obviously, time-domain equalization is more natural for SCM and frequency-domain correction, which generally is much easier to implement, is more natural for MCM. However, there is no theoretical objection to interchanging these techniques, since the signal can be shifted easily, although at some expense, from one domain to the other by means of the Fourier Transform.

A variant on the linear adaptive equalizer is the decision feedback equalizer (DFE) [33]. If an equalizer is operating so that the BER is low, then the channel frequency response is known fairly accurately. If so, the transmitted signal can be calculated at the receiver from the received signal and the known frequency response. The echo can then be calculated and the received signal perfectly corrected by subtracting the former from the latter. This method does not add noise as does a linear equalizer. However, to the extent that there are errors in the received signal, this process may increase the error rate. Simple reasoning suggests that there must be a threshold CNR above which the DFE improves the performance and below which it degrades the performance. The crucial situation is at threshold, where the question is whether a DFE extends or diminishes area coverage [40].

No frequency-domain DFE has been reported, but there seems to be no reason why this method could not be used in both systems, if it proved to extend the threshold.

2) Equalization of Dynamic Multipath: Rapidly changing echoes in the presence of a good deal of noise present a serious problem for linear equalizers, since it may not be possible to average over a time long enough to suppress noise in the calculation of equalizer parameters and at the same time follow the dynamic multipath. There seems to be little work reported on this issue. However, a recent paper dealing with MCM indicates that, if the moving echoes are sufficiently random, they may, indeed, be made to add constructively [34]. Presumably, if large fixed echoes could